Canad. Math. Bull. Vol. **17** (2), 1974.

ON THE DISTRIBUTION OF THE DIRECTION AND COLLINEARITY FACTORS IN DISCRIMINANT ANALYSIS

BY
R. P. GUPTA AND R. D. GUPTA

1. **Introduction.** If the samples of sizes $n_1, n_2, \ldots, n_{q+1}$ are available from $(q+1)$ normal populations with different mean vectors μ_α, $\alpha=1, 2, \ldots, q+1$ and the same covariance matrix Σ and if $x'=(x_1, x_2, \ldots, x_p)$ denotes the vector of p variates on which the measurements are made, then we obtain the following multivariate analysis of variance table.

Source	degrees of freedom (d.f.)	$p\times p$ matrices of sums of squares and sums of products (S.S. and S.P.)
Between groups	q	B_x
Within groups	$n-q$	W_x
Total	$n_1+n_2+\cdots+n_{q+1}-1=n$	B_x+W_x

We are interested in testing whether t hypothetical discriminant functions Γx, Γ $t\times p$ and of rank $t\le p$, are adequate for discrimination among $q+1$ groups. The hypothesis of goodness of fit of Γx consists of two aspects, (i) collinearity part, i.e., whether t functions Γx are adequate at all, and (ii) direction part, i.e., whether the vectors of Γ have preassigned components. For this purpose Bartlett [1], and Williams [4], propose the following criteria. They factorize Wilk's Λ criterion,

$$\Lambda = |W_x|/|B_x+W_x| \quad \text{as} \tag{1.1}$$
$$\Lambda = \Lambda_1\Lambda_2\Lambda_3$$

where

$$\begin{cases} \Lambda_1 = |\Gamma W_x\Gamma'|/|\Gamma(B_x+W_x)\Gamma'|, \\ \Lambda_2 = |B_x+W_x-B_x\Gamma'(\Gamma B_x\Gamma')^{-1}\Gamma B_x|/|B_x+W_x|\,\Lambda_1, \\ \Lambda_3 = \Lambda/\Lambda_1\Lambda_2. \end{cases} \tag{1.2}$$

The above authors also give an alternative factorization of Λ as

$$\Lambda = \Lambda_1\Lambda_4\Lambda_5 \tag{1.3}$$

where

$$\begin{cases} \Lambda_4 = |W_x+B_x\Gamma'(\Gamma B_x\Gamma')^{-1}\Gamma B_x|/|B_x+W_x|, \\ \Lambda_5 = \Lambda/\Lambda_1\Lambda_4. \end{cases} \tag{1.4}$$

https://doi.org/10.4153/CMB-1974-054-4 Published online by Cambridge University Press

Radcliffe [3] expressed Λ_1, Λ_2, Λ_3, Λ_4, and Λ_5 as functions of the elements of a triangular matrix T and then obtained their distributions. Radcliffe's method is very complicated and lengthy. Here we give a shorter and neater proof, which avoids all the transformations used by Radcliffe.

2. **Representation of Λ_2, Λ_3, Λ_4, Λ_5.** We make a linear transformation

$$z = Lx, \tag{2.1}$$

where L is $p\times p$ non-singular matrix such that $L\Sigma L'=I_p$. Under this transformation our discriminant functions are $\Gamma^* z$, where $\Gamma^*=\Gamma L^{-1}$. The "between" and "within" groups matrices for z are

$$B = LB_xL', \qquad W = LW_xL'. \tag{2.2}$$

Now we express Γ^* as

$$\Gamma^* = P[\theta:0]Q, \tag{2.3}$$

where P is a $t\times t$ orthogonal matrix, Q is a $p\times p$ orthogonal matrix, and θ^2 is the $t\times t$ diagonal matrix of the non-zero roots of $\Gamma^*\Gamma^{*\prime}$ or $\Gamma^{*\prime}\Gamma^*$. We further note that the densities of $C=QBQ'$, $D=QWQ'$ are identical to the densities of B, W respectively. Under the transformation (2.1) we have

$$\Lambda_1 = \frac{|\Gamma^* W\Gamma^{*\prime}|}{|\Gamma^*(B+W)\Gamma^{*\prime}|} = \frac{|D_1|}{|C_1+D_1|}, \tag{2.4}$$

where

$$C = \begin{bmatrix} C_1 & C_2 \\ C_2' & C_3 \end{bmatrix}\begin{matrix} t \\ p-t \end{matrix}, \qquad D = \begin{bmatrix} D_1 & D_2 \\ D_2' & D_3 \end{bmatrix}\begin{matrix} t \\ p-t \end{matrix}. \tag{2.5}$$
$$\quad t \quad p-t \qquad\qquad\qquad t \quad p-t$$

However, note that $|D_1|/|C_1+D_1|$ has the same distribution as that of $|W_1|/|B_1+W_1|$, where B and W are partitioned correspondingly to C and D. Thus there is no loss of generality in assuming $\Gamma^*=(I, 0)$, where I is a $t\times t$ identity matrix.

We define

$$\begin{cases} B_3^* = B_3-B_2B_1^{-1}B_2', \qquad W_3^* = W_3-W_2W_1^{-1}W_2', \\ S = B_2B_1^{-1}B_2'+W_2W_1^{-1}W_2'-(B_2+W_2)(B_1+W_1)^{-1}(B_2+W_2)'. \end{cases} \tag{2.6}$$

Using (2.2), $\Gamma^*=(I, 0)$, and (2.6) we immediately obtain

$$\Lambda_2 = |B_3^*+W_3^*|/|B_3^*+W_3^*+S|, \tag{2.7}$$

$$\Lambda_3 = |W_3^*|/|B_3^*+W_3^*|, \tag{2.8}$$

$$\Lambda_4 = |W_3^*+S|/|B_3^*+W_3^*+S|, \tag{2.9}$$

and

$$\Lambda_5 = |W_3^*|/|W_3^*+S|. \tag{2.10}$$

https://doi.org/10.4153/CMB-1974-054-4 Published online by Cambridge University Press

3. **Distributions of** Λ_2, Λ_3, Λ_4, Λ_5. If the given hypothesis of goodness of fit is true, B has a non-central Wishart distribution of rank t, with q d.f. and W has an independent (central) Wishart distribution with $(n-q)$ d.f. The density of B and W is

$$g(B, W) = K\,|B|^{(q-p-1)/2}|W|^{(n-p-q-1)/2}\exp\{-\tfrac{1}{2}\operatorname{tr}(B+W)\}\phi(B_1), \tag{3.1}$$

where $\phi(B_1)$ is the contribution to the density of B due to non-centrality of B. Here and elsewhere K (as a general symbol) denotes the normalizing constants of the densities. From (3.1), the joint density of B_1, B_2, B_3^*, W_1, W_2, W_3^* is seen to be

$$\begin{aligned} g(B_1, B_2, B_3^*, W_1, W_2, W_3^*) &= K\exp\{-\tfrac{1}{2}\operatorname{tr}(B_1+W_1+B_3^*+W_3^*+B_2B_1^{-1}B_2'+W_2W_1^{-1}W_2')\} \\ &\quad \times|B_1|^{(q-p-1)/2}\,|W_1|^{1/2(n-q-p-1)}\,|B_3^*|^{1/2(q-p-1)}\,|W_3^*|^{1/2(n-q-p-1)}\,\phi(B_1). \end{aligned} \tag{3.2}$$

Now in (3.1) we set $W_2=G_2-B_2$, G_2 is $p-t\times t$, and then we have

$$\begin{aligned} g(B_1, B_2, B_3^*, W_1, G_2, W_3^*) &= K\exp\{-\tfrac{1}{2}\operatorname{tr}[B_2(B_1^{-1}+W_1^{-1})B_2' \\ &\quad +G_2W_1^{-1}G_2'-G_2W_1^{-1}B_2'-B_2W_1^{-1}G_2'+B_1+W_1+B_3^*+W_3^*]\} \\ &\quad \times|B_1|^{1/2(q-p-1)}\,|B_3^*|^{1/2(q-p-1)}\,|W_1|^{1/2(n-q-p-1)}\,|W_3^*|^{1/2(n-q-p-1)}\,\Phi(B_1). \end{aligned} \tag{3.3}$$

By using the fact that $(B_1+W_1)^{-1}=W_1^{-1}-W_1^{-1}(B_1^{-1}+W_1^{-1})^{-1}W_1^{-1}$, we have

$$\begin{aligned} g(B_1, B_2, B_3^*, W_1, G_2, W_3^*) &= K\exp\{-\tfrac{1}{2}\operatorname{tr}[B_1+W_1+B_3^*+W_3^* \\ &\quad +(B_2-G_2W_1^{-1}(B_1^{-1}+W_1^{-1})^{-1})(B_1^{-1}+W_1^{-1})(B_2-GW_1^{-1}(B_1^{-1}+W_1^{-1})^{-1})']\} \\ &\quad \times\exp\{-\tfrac{1}{2}\operatorname{tr}G_2(B_1+W_1)^{-1}G_2'\}\,|B_1|^{1/2(q-p-1)}\,|B_3^*|^{1/2(q-p-1)} \\ &\quad \times|W_1|^{1/2(n-q-p-1)}\,|W_3^*|^{1/2(n-q-p-1)}\,\phi(B_1). \end{aligned} \tag{3.4}$$

Further setting

$$B_2 = T_2(B_1^{-1}+W_1^{-1})^{-1/2}+G_2W_1^{-1}(B_1^{-1}+W_1^{-1})^{-1}, \tag{3.5}$$

we find the joint density of B_1, T_2, B_3^*, W_1, G_2, W_3^* is

$$\begin{aligned} g(B_1, T_2, B_3^*, W_1, G_2, W_3^*) &= K\exp\{-\tfrac{1}{2}\operatorname{tr}(B_1+W_1+B_3^*+W_3^*+T_2T_2')\}\exp\{-\tfrac{1}{2}\operatorname{tr}G_2(B_1+W_1)^{-1}G_2'\} \\ &\quad \times|B_1|^{1/2(q-p-1)}\,|B_3^*|^{1/2(q-p-1)}\,|W_1|^{1/2(n-q-p-1)}\,|W_3^*|^{1/2(n-q-p-1)} \\ &\quad \times|B_1^{-1}+W_1^{-1}|^{(t-p)/2}\,\phi(B_1). \end{aligned} \tag{3.6}$$

https://doi.org/10.4153/CMB-1974-054-4 Published online by Cambridge University Press

By integrating over G_2, we find the joint density of B_1, T_2, B_3^*, W_1 and W_3^* to be

$$g(B_1, T_2, B_3^*, W_1, W_3^*)$$

$$= K \exp\{-\tfrac{1}{2}\operatorname{tr}(B_1+W_1)-\tfrac{1}{2}\operatorname{tr}(B_3^*+W_3^*+T_2T_2')\}\,|B_1|^{1/2(q-p-1)}\,|B_3^*|^{1/2(q-p-1)}$$

$$\times|W_1|^{(n-q-p-1)/2}\,|W_3^*|^{(n-q-p-1)/2}\,\phi(B_1)\,|B_1+W_1|^{(p-t)/2}\,|B_1^{-1}+W_1^{-1}|^{(t-p)/2}$$

$$= K \exp\{-\tfrac{1}{2}\operatorname{tr}(B_1+W_1)-\tfrac{1}{2}\operatorname{tr}(B_3^*+W_3^*+T_2T_2')\}\,|B_1|^{1/2(q-t-1)}\,|B_3^*|^{1/2(q-p-1)}$$

$$\text{(3.7)}\quad \times|W_1|^{(n-q-t-1)/2}\,|W_3^*|^{(n-q-p-1)/2}\,\phi(B_1).$$

From (3.7) it can be easily seen that B_1 and W_1 are independent of B_3^*, W_3^*, and T_2. We notice that $S=T_2T_2'$, i.e., S has a central Wishart distribution with t d.f. If $t<(p-t)$, then the density of S is pseudo Wishart with t d.f. Now we recall the following important result connected with Wishart matrices. If an $r\times r$ matrix A has a central Wishart distribution with f d.f. and if u_α, $\alpha=1, 2, \ldots, t$ are independent r component normal column vectors with $E(u_\alpha)=0$, $E(u_\alpha u_\alpha')=\Sigma$, where Σ is also the covariance matrix of the density of A. Then

$$\text{(3.8)}\qquad \Delta = |A|\Big/\Big|A+\sum_{\alpha=1}^{t} u_\alpha u_\alpha'\Big|,$$

is independently distributed of $(A+\sum_{\alpha=1}^{t} u_\alpha u_\alpha')$. The density of Δ is denoted by [Bartlett], 1 $\Lambda(f+t, r, t)$. Further, by writing

$$\text{(3.9)}\qquad \Lambda = \Lambda_1\Lambda_2\Lambda_3 = \frac{|W_1|}{|B_1+W_1|}\cdot\frac{|B_3^*+W_3^*|}{|B_3^*+W_3^*+T_2T_2'|}\cdot\frac{|W_3^*|}{|B_3^*+W_3^*|}$$

and using (3.7) we note that Λ_1 is independent of Λ_2 and Λ_3. Again Λ_3, by quoted property of Wishart matrices, is independent of $B_3^*+W_3^*$ and T_2, i.e. of Λ_2. It follows that Λ_1, Λ_2, Λ_3 are mutually independent.

Similarly, by writing

$$\text{(3.10)}\qquad \Lambda = \Lambda_1\Lambda_4\Lambda_5 = \frac{|W_1|}{|B_1+W_1|}\cdot\frac{|W_3^*+T_2T_2'|}{B_3^*+W_3^*+T_2T_2'|}\cdot\frac{|W_3^*|}{|W_3^*+T_2T_2'|},$$

and using (3.7) we note Λ_1 is independent of Λ_4 and Λ_5. However Λ_5 is independent of B_3^* and $W_3^*+T_2T_2'$, by the quoted property of Wishart matrices, and hence Λ_5 is independent of Λ_4. Thus Λ_1, Λ_4, Λ_5 are mutually independent.

From (3.7) it is easily observed that the densities of Λ_2, Λ_3, Λ_4, Λ_5 are respectively those of $\Lambda(n-t, p-t, t)$, $\Lambda(n-2t, p-t, q-t)$, $\Lambda(n-t, p-t, q-t)$, $\Lambda(n-q, p-t, t)$.

References

1. M. S. Bartlett, *The goodness of fit of a single hypothetical discriminant function in the case of several groups.* Ann. Eugen. **16** (1951) 199–214.

2. R. P. Gupta and D. G. Kabe, *Distribution of certain factors in discriminant analysis*. Ann. Inst. Statist. Math. **23** (1971) 97–103.

3. J. Radcliffe, *The distribution of certain factors occurring in discriminant analysis*. Proc. Camb. Phil. Soc. **64** (1968) 731–740.

4. Williams, E. J. *Some exact tests in multivariate analysis. Biometrika*. **39** (1952) 17–31.

DALHOUSIE UNIVERSITY,
HALIFAX, N. S., CANADA